

0590
1126

715

OIKE

RAW SEQUENCE LISTING

DATE: 12/07/2001

PATENT APPLICATION: US/09/845,335

TIME: 16:49:52

Input Set : N:\Crif3\RULE60\09845335.txt

Output Set: N:\CRF3\12072001\I845335.raw

3 <110> APPLICANT: CLOUGH, BARBARA
4 PREISER, PETER
5 WILSON, ROBERT
7 <120> TITLE OF INVENTION: AN EF-TU PROTEIN ENCODED ON THE PLASTID DNA OF THE
8 MALARIA PARASITE AND PROTEIN SYNTHESIS INHIBITORS
9 EFFECTIVE AS ANTI-MALARIAL COMPOUNDS
11 <130> FILE REFERENCE: N68837B GCW PJC DP
13 <140> CURRENT APPLICATION NUMBER: 09/845,335
14 <141> CURRENT FILING DATE: 2001-08-29
16 <150> PRIOR APPLICATION NUMBER: 09/140,466
17 <151> PRIOR FILING DATE: 1998-08-26
19 <150> PRIOR APPLICATION NUMBER: US 60/056,246
20 <151> PRIOR FILING DATE: 1997-08-28
22 <160> NUMBER OF SEQ ID NOS: 14
24 <170> SOFTWARE: PatentIn Ver. 2.1
26 <210> SEQ ID NO: 1
27 <211> LENGTH: 1230
28 <212> TYPE: DNA
29 <213> ORGANISM: Plasmodium falciparum
31 <220> FEATURE:
32 <221> NAME/KEY: CDS
33 <222> LOCATION: (1)..(1230)
35 <400> SEQUENCE: 1
36 atg aat aat aaa tta ttt tta aga aat aaa caa cat ata aat tta ggt 48
37 Met Asn Asn Lys Leu Phe Leu Arg Asn Lys Gln His Ile Asn Leu Gly
38 1 5 10 15
40 act ata ggg cat gta gat cat gga aaa act aca tta aca aca gct ata 96
41 Thr Ile Gly His Val Asp His Gly Lys Thr Thr Leu Thr Thr Ala Ile
42 20 25 30
44 tct tat tta tta aat tta caa gga tta tca aaa aaa tat aat tat tca 144
45 Ser Tyr Leu Leu Asn Leu Gln Gly Leu Ser Lys Lys Tyr Asn Tyr Ser
46 35 40 45
48 gat att gat tca gct cca gaa gaa aaa ata aga ggt att aca ata aat 192
49 Asp Ile Asp Ser Ala Pro Glu Glu Lys Ile Arg Gly Ile Thr Ile Asn
50 50 55 60
52 aca aca cat att gaa tat gaa act tta aca aaa cat tgt gct cat ata 240
53 Thr Thr His Ile Glu Tyr Glu Thr Leu Thr Lys His Cys Ala His Ile
54 65 70 75 80
56 gat tgt cca gga cat tcc gat tat att aaa aat atg att ata gga gcc 288
57 Asp Cys Pro Gly His Ser Asp Tyr Ile Lys Asn Met Ile Ile Gly Ala
58 85 90 95
60 aca caa atg gat ata gca att tta gta ata tct ata ata gat ggt ata 336
61 Thr Gln Met Asp Ile Ala Ile Leu Val Ile Ser Ile Ile Asp Gly Ile
62 100 105 110
64 atg cct caa act tat gaa cat tta tta tta ata aaa caa ata ggt ata 384
65 Met Pro Gln Thr Tyr Glu His Leu Leu Leu Ile Lys Gln Ile Gly Ile
66 115 120 125

ENTERED

RAW SEQUENCE LISTING

DATE: 12/07/2001

PATENT APPLICATION: US/09/845,335

TIME: 16:49:52

Input Set : N:\Crf3\RULE60\09845335.txt

Output Set: N:\CRF3\12072001\I845335.raw

```

68 aaa aat ata att att ttt tta aat aaa gaa gat tta tgt gat gat gtt 432
69 Lys Asn Ile Ile Ile Phe Leu Asn Lys Glu Asp Leu Cys Asp Asp Val
70 130 135 140
72 gaa tta ata gat ttt ata aaa tta gaa gta aat gaa tta tta att aaa 480
73 Glu Leu Ile Asp Phe Ile Lys Leu Glu Val Asn Glu Leu Leu Ile Lys
74 145 150 155 160
76 tat aat ttt gat tta aat tat ata cat ata tta act ggt tca gca tta 528
77 Tyr Asn Phe Asp Leu Asn Tyr Ile His Ile Leu Thr Gly Ser Ala Leu
78 165 170 175
80 aat gta ata aat ata att caa aaa aat aag gat tat gaa tta ata aaa 576
81 Asn Val Ile Asn Ile Ile Gln Lys Asn Lys Asp Tyr Glu Leu Ile Lys
82 180 185 190
84 tct aat att tgg ata caa aaa tta aat aat tta att caa ata att gat 624
85 Ser Asn Ile Trp Ile Gln Lys Leu Asn Asn Leu Ile Gln Ile Ile Asp
86 195 200 205
88 aat att ata ata cct act aga aaa att aat gat tac ttt tta atg tca 672
89 Asn Ile Ile Ile Pro Thr Arg Lys Ile Asn Asp Tyr Phe Leu Met Ser
90 210 215 220
92 ata gaa gat gta ttt tct ata aca ggt aga ggt aca gta gta aca ggt 720
93 Ile Glu Asp Val Phe Ser Ile Thr Gly Arg Gly Thr Val Val Thr Gly
94 225 230 235 240
96 aag att gaa caa gga tgt ata aat tta aat gat gaa att gaa att tta 768
97 Lys Ile Glu Gln Gly Cys Ile Asn Leu Asn Asp Glu Ile Glu Ile Leu
98 245 250 255
100 aaa ttt gaa aaa tca tct cct aat tta aca aca gtt ata gga tta gaa 816
101 Lys Phe Glu Lys Ser Ser Pro Asn Leu Thr Thr Val Ile Gly Leu Glu
102 260 265 270
104 atg ttt aaa aaa caa tta aca caa gca caa tcc gga gat aat gta ggt 864
105 Met Phe Lys Lys Gln Leu Thr Gln Ala Gln Ser Gly Asp Asn Val Gly
106 275 280 285
108 att tta tta aga aat att caa aaa aaa gat ata aaa aga ggt atg att 912
109 Ile Leu Leu Arg Asn Ile Gln Lys Lys Asp Ile Lys Arg Gly Met Ile
110 290 295 300
112 tta gca aca cct aat aaa tta aaa gta tat aag tct ttt ata gct gaa 960
113 Leu Ala Thr Pro Asn Lys Leu Lys Val Tyr Lys Ser Phe Ile Ala Glu
114 305 310 315 320
116 aca tat att tta act aaa gaa gaa ggt ggt cgt cat aaa cct ttt aat 1008
117 Thr Tyr Ile Leu Thr Lys Glu Glu Gly Gly Arg His Lys Pro Phe Asn
118 325 330 335
120 att gga tat aaa cct caa ttt ttt att cgt aca gta gat gtt act gga 1056
121 Ile Gly Tyr Lys Pro Gln Phe Phe Ile Arg Thr Val Asp Val Thr Gly
122 340 345 350
124 gaa att aaa aat ata tat tta aat gaa aat gta caa aaa gta gct ata 1104
125 Glu Ile Lys Asn Ile Tyr Leu Asn Glu Asn Val Gln Lys Val Ala Ile
126 355 360 365
128 cct gga gat aaa ata aca tta cat att gaa tta aaa cat tat ata gtg 1152
129 Pro Gly Asp Lys Ile Thr Leu His Ile Glu Leu Lys His Tyr Ile Val
130 370 375 380
132 ttg aca tta aat atg aaa ttt tct att aga gaa gga gga aaa aca ata 1200

```

RAW SEQUENCE LISTING

DATE: 12/07/2001

PATENT APPLICATION: US/09/845,335

TIME: 16:49:52

Input Set : N:\Crif3\RULE60\09845335.txt

Output Set: N:\CRF3\12072001\I845335.raw

```

133 Leu Thr Leu Asn Met Lys Phe Ser Ile Arg Glu Gly Gly Lys Thr Ile
134 385                      390                      395                      400
136 gga gca ggt att ata aca gaa ata aaa aat
137 Gly Ala Gly Ile Ile Thr Glu Ile Lys Asn
138                      405                      410
141 <210> SEQ ID NO: 2
142 <211> LENGTH: 410
143 <212> TYPE: PRT
144 <213> ORGANISM: Plasmodium falciparum
146 <400> SEQUENCE: 2
147 Met Asn Asn Lys Leu Phe Leu Arg Asn Lys Gln His Ile Asn Leu Gly
148 1                      5                      10                      15
150 Thr Ile Gly His Val Asp His Gly Lys Thr Thr Leu Thr Thr Ala Ile
151                      20                      25                      30
153 Ser Tyr Leu Leu Asn Leu Gln Gly Leu Ser Lys Lys Tyr Asn Tyr Ser
154                      35                      40                      45
156 Asp Ile Asp Ser Ala Pro Glu Glu Lys Ile Arg Gly Ile Thr Ile Asn
157                      50                      55                      60
159 Thr Thr His Ile Glu Tyr Glu Thr Leu Thr Lys His Cys Ala His Ile
160 65                      70                      75                      80
162 Asp Cys Pro Gly His Ser Asp Tyr Ile Lys Asn Met Ile Ile Gly Ala
163                      85                      90                      95
165 Thr Gln Met Asp Ile Ala Ile Leu Val Ile Ser Ile Ile Asp Gly Ile
166                      100                     105                     110
168 Met Pro Gln Thr Tyr Glu His Leu Leu Leu Ile Lys Gln Ile Gly Ile
169                      115                     120                     125
171 Lys Asn Ile Ile Ile Phe Leu Asn Lys Glu Asp Leu Cys Asp Asp Val
172                      130                     135                     140
174 Glu Leu Ile Asp Phe Ile Lys Leu Glu Val Asn Glu Leu Leu Ile Lys
175 145                     150                     155                     160
177 Tyr Asn Phe Asp Leu Asn Tyr Ile His Ile Leu Thr Gly Ser Ala Leu
178                      165                     170                     175
180 Asn Val Ile Asn Ile Ile Gln Lys Asn Lys Asp Tyr Glu Leu Ile Lys
181                      180                     185                     190
183 Ser Asn Ile Trp Ile Gln Lys Leu Asn Asn Leu Ile Gln Ile Ile Asp
184                      195                     200                     205
186 Asn Ile Ile Ile Pro Thr Arg Lys Ile Asn Asp Tyr Phe Leu Met Ser
187                      210                     215                     220
189 Ile Glu Asp Val Phe Ser Ile Thr Gly Arg Gly Thr Val Val Thr Gly
190 225                     230                     235                     240
192 Lys Ile Glu Gln Gly Cys Ile Asn Leu Asn Asp Glu Ile Glu Ile Leu
193                      245                     250                     255
195 Lys Phe Glu Lys Ser Ser Pro Asn Leu Thr Thr Val Ile Gly Leu Glu
196                      260                     265                     270
198 Met Phe Lys Lys Gln Leu Thr Gln Ala Gln Ser Gly Asp Asn Val Gly
199                      275                     280                     285
201 Ile Leu Leu Arg Asn Ile Gln Lys Lys Asp Ile Lys Arg Gly Met Ile
202                      290                     295                     300
204 Leu Ala Thr Pro Asn Lys Leu Lys Val Tyr Lys Ser Phe Ile Ala Glu

```

RAW SEQUENCE LISTING

DATE: 12/07/2001

PATENT APPLICATION: US/09/845,335

TIME: 16:49:52

Input Set : N:\Crif3\RULE60\09845335.txt

Output Set: N:\CRF3\12072001\I845335.raw

```

205 305          310          315          320
207 Thr Tyr Ile Leu Thr Lys Glu Glu Gly Gly Arg His Lys Pro Phe Asn
208          325          330          335
210 Ile Gly Tyr Lys Pro Gln Phe Phe Ile Arg Thr Val Asp Val Thr Gly
211          340          345          350
213 Glu Ile Lys Asn Ile Tyr Leu Asn Glu Asn Val Gln Lys Val Ala Ile
214          355          360          365
216 Pro Gly Asp Lys Ile Thr Leu His Ile Glu Leu Lys His Tyr Ile Val
217          370          375          380
219 Leu Thr Leu Asn Met Lys Phe Ser Ile Arg Glu Gly Gly Lys Thr Ile
220 385          390          395          400
222 Gly Ala Gly Ile Ile Thr Glu Ile Lys Asn
223          405          410
227 <210> SEQ ID NO: 3
228 <211> LENGTH: 409
229 <212> TYPE: PRT
230 <213> ORGANISM: Anacystis nidulans
232 <400> SEQUENCE: 3
233 Met Ala Arg Ala Lys Phe Glu Arg Thr Lys Pro His Ala Asn Ile Gly
234 1          5          10          15
236 Thr Ile Gly His Val Asp His Gly Lys Thr Thr Leu Thr Ala Ala Ile
237          20          25          30
239 Thr Thr Val Leu Ala Lys Ala Gly Met Ala Lys Ala Arg Ala Tyr Ala
240          35          40          45
242 Asp Ile Asp Ala Ala Pro Glu Glu Lys Ala Arg Gly Ile Thr Ile Asn
243          50          55          60
245 Thr Ala His Val Glu Tyr Glu Thr Gly His Arg His Tyr Ala His Val
246 65          70          75          80
248 Asp Cys Pro Gly His Ala Asp Tyr Val Lys Asn Met Ile Thr Gly Ala
249          85          90          95
251 Ala Gln Met Asp Gly Ala Ile Leu Val Val Ser Ala Ala Asp Gly Pro
252          100          105          110
254 Met Pro Gln Thr Arg Glu His Ile Leu Leu Ala Lys Gln Val Gly Val
255          115          120          125
257 Pro Asn Ile Val Val Phe Leu Asn Lys Glu Asp Met Val Asp Asp Ala
258          130          135          140
260 Glu Leu Leu Glu Leu Val Glu Leu Glu Val Arg Glu Leu Leu Ser Ser
261 145          150          155          160
263 Tyr Asp Phe Pro Gly Asp Asp Ile Pro Ile Val Ala Gly Ser Ala Leu
264          165          170          175
266 Gln Ala Leu Glu Ala Ile Gln Gly Gly Ala Ser Gly Gln Lys Gly Asp
267          180          185          190
269 Asn Pro Trp Val Asp Lys Ile Leu Lys Leu Met Glu Glu Val Asp Ala
270          195          200          205
272 Tyr Ile Pro Thr Pro Glu Arg Glu Val Asp Arg Pro Phe Leu Met Ala
273          210          215          220
275 Val Glu Asp Val Phe Thr Ile Thr Gly Arg Gly Thr Val Ala Thr Gly
276 225          230          235          240
278 Arg Ile Glu Arg Gly Ser Val Lys Val Gly Glu Thr Ile Glu Ile Val

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/845,335

DATE: 12/07/2001

TIME: 16:49:52

Input Set : N:\Crif3\RULE60\09845335.txt

Output Set: N:\CRF3\12072001\I845335.raw

```

279          245          250          255
281 Gly Leu Arg Asp Thr Arg Ser Thr Thr Val Thr Gly Val Glu Met Phe
282          260          265          270
284 Gln Lys Thr Leu Asp Glu Gly Leu Ala Gly Asp Asn Val Gly Leu Leu
285          275          280          285
287 Leu Arg Gly Ile Gln Lys Thr Asp Ile Glu Arg Gly Met Val Leu Ala
288          290          295          300
290 Lys Pro Gly Ser Ile Thr Pro His Thr Lys Phe Glu Ser Glu Val Tyr
291 305          310          315          320
293 Val Leu Lys Lys Glu Gly Gly Arg His Thr Pro Phe Phe Pro Gly
294          325          330          335
296 Tyr Arg Pro Gln Phe Tyr Val Arg Thr Thr Asp Val Thr Gly Ala Ile
297          340          345          350
299 Ser Asp Phe Thr Ala Asp Asp Gly Ser Ala Ala Glu Met Val Ile Pro
300          355          360          365
302 Gly Asp Arg Ile Lys Met Thr Val Glu Leu Ile Asn Pro Ile Ala Ile
303          370          375          380
305 Glu Gln Gly Met Arg Phe Ala Ile Arg Glu Gly Gly Arg Thr Ile Gly
306 385          390          395          400
308 Ala Gly Val Val Ser Lys Ile Leu Gln
309          405
313 <210> SEQ ID NO: 4
314 <211> LENGTH: 408
315 <212> TYPE: PRT
316 <213> ORGANISM: Cryptomonas phi
318 <400> SEQUENCE: 4
319 Met Ala Arg Asp Lys Phe Glu Arg Ser Lys Pro His Val Asn Ile Gly
320 1          5          10          15
322 Thr Ile Gly His Val Asp His Gly Lys Thr Thr Leu Thr Ala Ala Ile
323          20          25          30
325 Ser Ala Thr Leu Ser Gln Tyr Thr Gly Lys Ser Lys Lys Phe Asp Glu
326          35          40          45
328 Ile Asp Ser Ala Pro Glu Glu Lys Ala Arg Gly Ile Thr Ile Asn Thr
329          50          55          60
331 Ala His Val Glu Tyr Glu Thr Asp Lys Trp Tyr Tyr Ala His Val Asp
332 65          70          75          80
334 Cys Pro Gly His Ala Asp Tyr Val Lys Asn Met Ile Thr Gly Ala Ala
335          85          90          95
337 Gln Met Asp Gly Ala Ile Leu Val Cys Ser Ala Ala Asn Gly Pro Met
338          100          105          110
340 Pro Gln Thr Arg Glu His Ile Leu Leu Ala Lys Gln Val Gly Val Pro
341          115          120          125
343 Tyr Ile Val Val Phe Leu Asn Lys Ala Asp Met Val Asp Asp Glu Glu
344          130          135          140
346 Leu Leu Glu Leu Val Gln Leu Glu Val Gln Glu Leu Leu Glu Lys Tyr
347 145          150          155          160
349 Asp Phe Pro Gly Ser Glu Ile Pro Phe Val Ala Gly Ser Ala Leu Leu
350          165          170          175
352 Ala Leu Glu Ala Val Ala Asn Asn Pro Thr Ile Lys Arg Gly Glu Asp

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/845,335

DATE: 12/07/2001

TIME: 16:49:53

Input Set : N:\Crf3\RULE60\09845335.txt

Output Set: N:\CRF3\12072001\I845335.raw